Sive Finlay: Thesis Corrections

Please find a copy of the examiners’ comments from Nicola Marples (NM) and Adam Algar (AA) (italics) and my responses below. Responses pre-fixed by SF are my comments, responses with no pre-fix are copies of the new text within the thesis. I have divided these into my responses to general comments and specific comments (references to particular pages and line numbers).

**General comments**

*NM: The weakest part of the thesis was the discussion, and I felt that the interesting differences between the expected level of diversity in the tenrecs and the results found were not adequately explored. The conclusion was that tenrecs do not show more diversity than the moles but most of the results showed that they did, just not to the extent expected. I think this should be teased out and discussed a little more thoroughly, and have marked on the script where this could be done. In addition I felt that the explanation of some of the more technical parts of the method was a little lacking in detail, especially around semi-landmarks and Procrustes distances, which are not terms generally known to the non-specialist.*

*AA: Although the dissertation is excellent in its aims, writing, and in the thoroughness of its methods, it is let down a bit by the interpretation and discussion of the results. I think minor corrections need to be resolved before the thesis can be accepted. Most importantly, I am not convinced that Sive’s main conclusion that tenrecs are not more morphologically diverse than golden moles is consistent with her data. Most tests actually showed tenrecs to be more diverse than golden moles. When 31 tenrec species were included, 2/4 traits showed more morphological diversity in tenrecs using t-tests and 4/4 when permutation tests were used. Excluding most Microgale species resulted in 3/4 significant t-tests and 4/4 significant permutation tests. Sive expresses scepticism regarding the lateral skull and mandible results but why these should be discounted was not fully justified, in my mind. Certainly, while the evidence are mixed, there seems to be more evidence suggesting that tenrecs are more morphologically diverse, especially when the over-representation of Microgale is accounted for. I think Sive needs to reconsider her conclusion that overall there is no evidence for exceptional morphological diversity in tenrecs. The opposite conclusion seems to be the case, especially given the permutation results, which to me are more convincing than the parametric tests.*

*I also found the Discussion to be a bit thin and would like to see a consideration of how the phylogenetic branching pattern within the tenrec and golden mole clade could have affected disparity comparisons between the two sister clades (e.g. under Brownian motion), though I doubt there is a sufficiently well resolved phylogeny to test this directly (and, regardless, there may not be a sufficient sample size of golden moles available for a meaningful test). Also, while Sive accounted for differences in sample size using a randomization test, she did not discuss the implications of lower and, I assume, non-random, sampling (completeness) of the golden mole clade.*

**Specific comments**

**Chapter 1: Introduction**

*p10, L175 NM: Just said that in previous paragraph*

SF: I cut the repetition here (the reference to the Slice 2007 paper) and incorporate the information into the previous paragraph:

“Furthermore, linear measurements of morphological traits can restrict our understanding of overall morphological variation: morphometric studies based on caliper measurements of particular features often fail to capture the overall shape of a specific structure (Slice2007). A distance matrix of measurements between specific points is unlikely to give a completely accurate representation of a three dimensional structure (Rohlf1993).”

*p10, L180 NM: Still only two-dimensional though - how do you get around that?*

SF: Please see the additional paragraph below the final introduction comment

*p11, L212 NM: Swap between adaptive radiation and convergence repeatedly and in the same sentence but these are different processes independent of each other. You can radiate without converging on an already existing form. So we need a little more distinction between these processes.*

SF: I tried to make this distinction clearer throughout the introduction:

“Therefore, it appears that tenrecs represent an adaptive radiation of species which filled otherwise vacant ecological niches **and, in doing so, developed convergent similarities with other small mammal species** (Soarimalala2011).”

“**Adaptive radiations and convergent evolution are separate evolutionary processes yet morphological diversity is common to both:** morphological diversity is an important **(though not defining, Olson2009)** feature of adaptive radiations (Losos2010a and it also informs our understanding of convergent phenotypes (Muschick2012).”

*p12, L219 NM: these data*

Fixed

*p12, (end) NM: Intro a bit thin on detail of other work. Also is geometric morphometrics the only possible method available? If not, tell us about them and why you chose to use this method. Also not convincing about what exactly one needs to know and can’t gain from single or linear measurements.*

SF: I’ve added the following paragraphs which discuss the relative merits of different morphometric approaches

“Comparing shape variation using single, linear measurements may be appropriate in some cases (e.g. Morales2013}) but the issues highlighted above are important limitations to consider. Geometric morphometric approaches help to overcome some of the issues associated with traditional morphological studies by using a system of Cartesian landmark coordinates to define anatomical points (Adams2004).

Geometric morphometric techniques capture more of the true, overall anatomical shape of particular structures (Mitteroecker2009). There are two main approaches to geometric morphometrics: two dimensional (2D) and three dimensional studies (3D). Three dimensional studies use 3D-scanners to create complete, virtual representations of anatomical features. However, the cost and time involved in scanning objects makes this approach impractical for large, cross-taxa comparisons of shape variation. In contrast, 2D geometric morphometric studies are commonly used to analyse three-dimensional morphological shape and are appropriate for cross-species comparisons (e.g. Muschick2012, Panchetti2008, Wroe2007, Marcus2000).

Of course there are potential biases associated with 2D morphometric approaches as they rely on 2D landmarks to represent 3D anatomical features. However, sensitivity analyses have demonstrated that bias from 2D representation of 3D structures is unlikely to be a significant issue for interspecific studies as the overall shape variation among species is greater than discrepancies introduced by using 2D morphometric techniques (Cardini2014). In addition, using multiple 2D approaches to analyse 3D shape (e.g. comparing skull shape using dorsal, ventral and lateral views) generates more robust shape variation results instead of summarising 3D anatomical shape using just one set of 2D photographs (Arnqvist1998).”

**Chapter 2: Data collection and processing**

*p17, L328 NM: full name if this is the first use of this abbreviation*

“… thin plate spline (TPS) software…”

*p17, L333 NM: Explain what these are [landmarks and semilandmarks]*

“I digitised landmarks (discrete points) and semilandmarks (points on a geometric feature, such as a curve, which are defined by their relative position on that feature, Zelditch2012} on every image individually.”

*p17, L339 NM: How did you decide which features on the skull to measure?*

“I selected my landmarks based on their relevance to the overall anatomical shape of the specimen and the ease with which they could be reliably identified on different species.”

*p17, L349 NM: Do you mean the length of the whole outline?*

SF: Yes, I changed it to read “I calculated the length of each re-sampled curve as a percentage of the total length of the outline”

*p17, L355 NM: This just gratuitously adds a 5% error in the shape measurement though… why not choose the smallest number which gave 100% the same shape?*

SF: A 5% error rate is usually acceptable and I think it’s reasonable given that it could be argued that my choices are already biased towards having many semilandmarks without many landmarks. I added the sentence below to justify my choice:

“This 95% cut off acts as a balance between accurate representation of the curve and avoiding unnecessary over-sampling (MacLeod2012).”

*p18, L365 NM: were*

SF: Fixed

*p18, L367 NM: selected how?*

SF: see additional sentence added on p17, L336

“I selected my landmarks based on their relevance to the overall anatomical shape of the specimen and the ease with which they could be reliably identified on different species.”

*p22, L395 NM: full name first time*

…thin plate spline utilities software (TPSUtil)…

*p22, L401 NM: Need to tell us more about this: I went and looked it up and now understand what you did, but I couldn’t from this description. Explain what a procrustes superimposition is and what it is designed to do, and what a procrustes distance is which you are trying to minimise.*

*p22, L400 NM: I don’t understand what sliding the semilandmarks means from this explanation. Please explain further both how you do it and what it achieves.*

“Procrustes superimposition creates a common set of Procrustes-superimposed landmarks that can be used to compare differences in shape that are independent of differences in size, location and orientation of the specimens (Webster2010). It is necessary to slide semilandmarks because equidistant points along a curve do not necessarily correspond to geometric or biological homology across specimens (Gunz2013). Sliding semilandmarks along a curve during Procrustes superimposition removes this effect of arbitrary semilandmark spacing by optimising the position of the semilandmarks with respect to the average shape of the entire sample (the average of all the Procrustes shape coordinates). Sliding semilandmarks by minimising Procrustes distance (the sum of squared distances between corresponding points of two Procrustes-superimposed shapes, Zelditch2012) allows each semilandmark to slide separately without being influence by the position of other landmarks or semilandmarks (Gunz2013).”

*p24, L457 NM: Not sure this results paragraph should be in the methods chapter and it means you never present theses data. Should put it at the start of the results chapter in a section of its own showing validation of the method and presenting the data for the validation.*

SF: I think it is appropriate to keep this paragraph here because the results relate to methodological error checking rather than any results that correspond to my diversity analyses. This section is validating the overall approach of using geometric morphometrics to assess shape variation. I think that referring back to it from the results chapter would be clumsy and wouldn’t fit with the rest of chapter 3. I’ve quoted that both photo identity and image replicate explained less than 0.0001% of the overall variation: I don’t think presenting the exact data behind that figure is necessary (and it would possibly look quite messy).

**Chapter 3**

p27, L503 NM: Are golden moles nice and evenly distributed between genera or do they need correcting for a populous genus like you did for the Microgale in the tenrecs?

*p27, L509 NM: Why 95%? Do others use this figure? If so, reference them.*

“Selecting the number of PC axes based on their cumulative variance (e.g. Brusatte2008, Collar2006) rather than through an arbitrary cut off standardised my diversity comparisons across the different analyses.”

*p29, L533 NM: How given that you’re quoting a mean?*

“…overall diversity in tenrecs could be higher if diversity is simply a function of sample size.”

*p29, L534 NM: Not enough explanation to say you did it and it accounted for the problem. Tell us how you do this and what it achieves exactly. Why not just do Euclidean distance per species to get rid of the problem?*

SF: I’m sorry for not making it clear that the following paragraph outlines detailed steps for how my permutation tests account for the potential issue of differences in sample size. I’ve added a lead in (“As described in the paragraph below,”) to make it clearer.

I’m not sure what you mean by Euclidean distance per species, but the permutation tests do account for differences in sample size (“…the 1000 permuted values of differences in morphological diversity create a distribution of the expected difference in diversity between a group of sample size 31 (or 17 in the case of the subsetted tenrec data) compared to a group of sample size 12 under the null hypothesis that the two groups have the same morphological diversity.”)

**Chapter 4**

*p36, L633 NM: Sounds like the original data set didn’t have Microgales in it. Rephrase to say when you excluded the over-representation of that family.*

“This appears to be the case for my data: it was only when I reduced my data set to include a sub-sample of the *Microgale* tenrecs instead of the entire Genus that I found higher morphological diversity in tenrecs compared to golden moles…”

p36, L641 NM: So despite the fact that the more trustworthy analysis, when you corrected for the over-representation of the Microgales, gave you a higher diversity in all 3 measures, you still conclude overall that tenrecs aren’t more diverse? This seems a bit perverse to me. Why do the correction if you don’t believe in its results? You need to justify doing this!

*p37, L622 NM: What do they use these structures for? Does that give you a clue as to why they are different?*

These areas are the main attachment sites for mandibular muscles but, given that golden moles have generalised diets and dig with their limbs rather than their jaws (Bronner1995), there does not appear to be an obvious explanation for why these structures should be more diverse in golden moles than in tenrecs.

*p37, L675 NM: But you checked your methodology accuracy - refer here to that and be brave enough to suggest it’s of biological importance!*

SF: I re-phrased this paragraph conclusion to strengthen my position.

“I was careful to ensure that the methods I used were accurate and free from methodological bias or statistical artefacts (see chapter 2). However, alternative approaches with landmarks that emphasise shape variation in different parts of the skulls and mandibles could produce different results. Further investigation is required to determine whether apparently high morphological diversity in golden mole mandibles is a consistent outcome from alternative approaches to shape analysis.”

*p37, L676 NM: Or should this be called future work?*

I think Caveats is appropriate because the main point of the section is to highlight what I feel are the shortcomings of my approach. I include brief mentions of possible future work to show that these obstacles are not insurmountable.

p38, L697 NM: I’d like to see a discussion paragraph somewhere of the alternative ways of looking at diversity and their pros and cons.

p38, L702 NM: Ok, with that caveat it’s okay but we need a discussion somewhere about whether it’s legitimate to take the sub-set as the result, or whether you have to take the whole family as the overall result.

p38, L710 NM: Is required before we can do what? Say tenrecs aren’t more diverse than other groups? Won’t other features just add to the confusion about what we mean by diverse?

**Appendices**

*p46 (table) NM: I’m not sure how useful it is to have the descriptions of what measurements were taken without the measurements. Why not put both bits in the same place?*

I think this section may have been confused. This table describes the linear (caliper) measurements I took of the skulls and limbs. It does not relate to any of my landmark measurements. I did not use of the linear measurements data for my thesis because it was extra data that would have been used for a PhD. I mentioned the measurements here to show that they exist and that they are available on Figshare: the measurement tables have over 25,000 rows each so it would be impractical to list them all here.

*p50 (end of page) NM: With what results?*

I didn’t include the error checking results because the linear measurements are not part of my thesis. The reason for describing my method here was for future repeatability: if anyone wants to download and use my measurement data they can see how it was cleaned for errors. The data that I’ve uploaded to Figshare only includes the cleaned (post-error checking) measurements.